Insurance: Mathematics and Economics 99 (2021) 294–308

Contents lists available at ScienceDirect

Insurance: Mathematics and Economics

journal homepage: www.elsevier.com/locate/ime

Cause-specific mortality rates: Common trends and differences

Séverine Arnold, Viktoriya Glushko *

Department of Actuarial Science, Faculty of Business and Economics (HEC Lausanne), University of Lausanne, Switzerland

ARTICLE INFO

Article history:
Available online 15 April 2021

Keywords:
Causes of death
Dependence
Cointegration
Common stochastic factors
Aging

ABSTRACT

In contrast to the traditional approach that uses total mortality rates, we want to gain additional insight into the past development of mortality by concentrating on a more detailed breakdown of mortality data, namely by causes of death. We work with the data from five developed countries (USA, Japan, France, England and Wales, and Australia), two sexes, and split the mortality rates into five main groups of causes of death (Infectious&Parasitic, Cancer, Circulatory diseases, Respiratory diseases, and External causes). As it was shown in Arnold and Sherris (2016), these time series of cause-specific mortality rates are cointegrated and so, there exist long-run equilibrium relationships between them. While the previous research focused on the stationary part of the system of cause-specific mortality rates, in the present paper we study its non-stationary part. For this we explicitly extract common stochastic trends from the original variables and compare them across the different datasets. By testing cointegration assumptions about these trends, we are able to get a better representation and understanding of how cause-specific death rates are evolving. We believe that common patterns emerging from such analysis could indicate a link to more fundamental biological processes such as aging.

1. Introduction

The ever decreasing mortality rates represent one of the biggest challenges that the insurance industry has ever faced. It is hence very important to understand well the past developments of mortality and be able to build sustainable forecasts for the future. This colossal task has been occupying many researchers as well as practitioners for several decades and resulted in innumerable number of models, approaches, and practices. An interested reader can find a review thereof in Booth and Tickle (2008), Cairns (2013) and Debón et al. (2006), including their references. While the most of the existing models deal with the all-cause mortality, we believe that integrating the information on different causes of death into the model can bring additional insight and improve the model's fit. In the same time, cause-specific mortality rates are dependent, but the dependency between them is inobservable and so, highly difficult to model.

Cointegration analysis represents a tool that allows us to take into account the dependency between the non-stationary variables. Two or more variables are said to be cointegrated if there exists a linear combination of them that is stationary. Such linear combination describes then the link (or the dependency) between the variables in the long run and can be included into a vector autoregressive model. This approach was initially developed to model the econometric time series, but later also gained popularity in the field of mortality modeling. As a vector of age-specific mortality rates usually has more elements than a cointegration relation can incorporate, some authors overcame this difficulty by concentrating on the pairwise cointegration between the age-specific mortality rates (Darkiewicz and Hoedemakers, 2004) or cointegration in a subset of mortality rates, e.g. higher ages (Lazar and Denuit, 2009). In Gaille and Sherris (2011) the authors reduced the age dimension by applying cointegration analysis to the parameters of the Heligman-Pollard model. Using total mortality rates, Njenga and Sherris (2011) were able to formulate a cointegrated model incorporating five country-specific mortality rates, whereas Arnold and Sherris (2013) applied a model allowing for cointegration relations between five cause-specific mortality rates. Also, very productive was the idea to apply the cointegration analysis to the time trends extracted from the mortality rates by the means of the Lee–Carter model. In this way, cointegration relations between the population-specific time trends were used in modeling and forecasting mortality rates by Li and Hardy (2011), Yang and Wang (2013), Zhou et al. (2014), and Hunt and Blake (2015) in a multi-population setting. Salhi and Loisel (2017) also studied the cointegration relations between two populations, but used the pairwise cointegration between the age-specific mortality rates for higher ages. A further example is the work by Li and Lu (2017) who enrich the vector-autoregression model for age-specific mortality rates with a cointegration element that ensures the non-divergence of the mortality rates at different ages.

* Corresponding author.
E-mail addresses: severine.arnold@unil.ch (S. Arnold), viktoriya.glushko@unil.ch (V. Glushko).

https://doi.org/10.1016/j.insmatheco.2021.03.027

For our part, we want to concentrate on the different causes of death with the help of the cointegration analysis. In this relation, Arnold and Sherris (2013) showed that the cause-specific mortality rates were cointegrated and confirmed this finding for both sexes in ten different countries. Arnold and Sherris (2016) identified the optimal model structure for the cause-specific mortality rates that allowed the authors to spot the difference in behavior between the exo and endogenous causes of death in the long run. Using a slightly different set of optimal models, Arnold and Glushko (2021) analyzed the short- and long-term interactions between the cause-specific mortality rates, namely how each cause-specific mortality rate impacts and reacts to the shocks received from the rest of the causes. To achieve this goal, Arnold and Glushko (2021) analyzed the *stationary* part of the model by studying the short run (autoregressive) and the long run (cointegration) terms.

Now we want to take a different angle: if certain variables are cointegrated, they must share common stochastic trends that are eliminated in the cointegration relation. In other words, as shown by Stock and Watson (1988), for any cointegrated variables x_t and y_t there must exist a common factor representation of the following form:

$$\begin{bmatrix} y_t \\ x_t \end{bmatrix} = \begin{bmatrix} A \\ 1 \end{bmatrix} f_t + \begin{bmatrix} \tilde{y}_t \\ \tilde{x}_t \end{bmatrix}, \tag{1}$$

where $\tilde{y}_t$ and $\tilde{x}_t$ are both I(0), $(1, -A)$ is the cointegrating vector and f_t represent the common stochastic factors impacting the system. In the present paper, we will study the common stochastic trends shared by the cause-specific mortality rates, i.e. the *non-stationary* part of the model which is usually ignored when only cointegration relations are analyzed (as this was done in Arnold and Sherris (2016) and Arnold and Glushko (2021)). To achieve this, we will, first, recover these common factors, and second, see if any similarities or common patterns can be found between different countries. For the sake of comparability and consistence, we use the same set of models as in Arnold and Glushko (2021).

In Arnold and Sherris (2016) the authors found that the cointegration techniques applied to the cause-specific mortality data provided a first bridge between econometrics and biology, two areas of study essential for actuaries, and that cointegration relations reflected the biological theory on aging. This became possible once the distinction between the exo and endogenous causes, developed by biologists and demographers, was considered. Although this classification is not clear-cut, to the first group of causes of death most researchers attribute various external or environmental factors that produce death, while the endogenous causes of death correspond to biological forces that lead to death (Carnes et al., 2006). From the distinction between the exo and endogenous causes follows the idea that endogenous mortality reflects fundamental processes of the human body referred to as the biological processes of aging. Since the authors found that only the endogenous causes appeared in the cointegration relations, *these relations may have the potential to capture the statistical characteristics of the biological processes of aging* (Arnold and Sherris, 2016). Also, common stochastic trends shared by the cause-specific mortality rates could represent the aging processes, because the aging process is known to be stochastic (Hayflick, 2004) and a potential mixture of several stochastic processes (Holliday, 2004). Then, the biological aging of the body is indeed the underlying risk factor influencing the causes of death (Olshansky et al., 2002) and is captured by the common stochastic trends of the cointegrating system. However, Arnold and Sherris (2016) only analyzed the cointegration relations and did not try to find an expression for these common stochastic trends.

In the present work, we want to go further and investigate the first intuition of Arnold and Sherris (2016) by recovering and studying the common stochastic factors f_t as they are shown in (1). Gonzalo and Granger (1995) mention several reasons why it may be interesting to recover f_t, namely to (1) simplify a complex model; (2) decompose (y_t, x_t) into two components $(f_t, (\tilde{y}_t, \tilde{x}_t))$ that transmit different kinds of information (while the permanent component f_t corresponds to the trends present in the data, the transitory component $(\tilde{y}_t, \tilde{x}_t)$ conveys the information on the short-term shocks and cycles); (3) study the subdivisions of a large system by first finding the common factors in every subdivision and then studying the cointegration among them.

Cause-specific mortality rates are a reflection of numerous impacts and processes that range from medical advances, changes in lifestyles and nutrition to epidemics and aging. In spite of the evident differences between the past experience of different countries, it is still reasonable to expect that some of these processes will be present in all countries due to their universal character, e.g., aging. So our objective is to recover the common stochastic factors shared by the cause-specific mortality rates from one country and propose an approach allowing us to make comparisons with a goal to find similarities across five countries. Should a certain pattern be found in all tested countries, this would allow us to expect that we are dealing with some fundamental process common to human species as such. So in this way, we are able to shed light on the processes that underlie the development of mortality in every tested country. Although we are not yet able to identify these process with a certainty, we believe that the possibility to give them a mathematical expression can help to improve our understanding of the past development of the mortality rates.

To achieve this, we, first, construct the set of common factors in every country that has a lower number of dimensions than the initial variables. As it turns out that this number is still too high to allow direct comparison between the countries, we further concentrate the information available in the set by using the principal component analysis. When comparing the charts of the principal components, we noticed that the form of the first elements was similar on all charts. To study further the observed resemblance, we test for cointegration using the Johansen maximum likelihood tests (Johansen, 1988). This allows us to examine cointegration in a large system of data variables (5 cause-specific mortality rates for 5 countries and 2 sexes) that would not have been possible without the initial reduction in its dimensionality. At the next step, we find that once we put together the first components extracted in every country, they are indeed cointegrated.

We believe that, although the cause-specific mortality rates show different development patterns across countries, this observation could mean that the common factors reflect some similar intrinsic stochastic processes which occur in every dataset (country). As we work with the cause-specific mortality rates, these processes could point to some fundamental mechanisms, typical for human species, such as biological aging.

The paper is organized as follows: in Section 2 we briefly present the data that we used in the study, then continue with some theoretical notions of the cointegration analysis and lay out the methodology used to extract and condense the common stochastic factor in Section 3. The application of these tools to the data is presented in Section 4, while Section 5 concludes the paper.

2. Data

For this study, we used the same data as in Arnold and Glushko (2021) and refer the interested reader to this article for the details on the data preparation process. We mention here the main points.

- Data were retrieved from the WHO Mortality Database (World Health Organization, 2016) that contains the mid-year population and the death counts by country, year, sex, age group, and cause of death. The earliest observations in this database go as far back as 1950.
- We considered the following countries and observation periods: USA (1950–2007), Japan (1950–2013), France (1952–2011), England and Wales (1950–2013), and Australia (1950–2004), subsequently shortened to US, JP, FR, E&W, and AU respectively. These countries were chosen as they participate in the database from the onset, belong to the developed countries with important population sites and are located in different parts of the world (North America, Aisa, Europe, and Oceania). This choice ensures that we have at our disposal the longest possible series of rich and reliable observations.
- WHO defines the causes of death according to the International Classification of Diseases (ICD). By applying the comparability ratios[1] we ensured that observations were comparable across the different versions of ICD that switched from ICD-7 to ICD-10 since the inception of the database.
- Causes of deaths were split in groups of infectious and parasitic diseases, cancer, diseases of the circulatory system, diseases of the respiratory system, and external causes. These are the most important groups of causes of deaths. They account for approximately 70%–80% of deaths in recent years and made up approximately 50%–70% of deaths at the onset of the observations.
- Central death (or mortality) rates were calculated as the number of deaths by age, sex, and cause divided by the mid-year population by age and sex.
- Central death rates were age-standardized using the US male population of 2007 as the standard population (more details on this procedure are given in Appendix A). We will work with the total rates and will not differentiate by age, otherwise there would be more variables than the cointegration analysis can accommodate. Age-standardized death rates for selected years using the US males population base as well as the charts showing their evolution over the entire observation period are shown in Appendix A, Table A.1 and Fig. A.1.
- Equations were estimated for the time series of ordered $(n \times 1)$ vectors of the cause-specific mortality rates after the application of the natural logarithm:

$$\mathbf{y}_{t,s,c} = \begin{pmatrix} log(m^{I\&P}_{t,s,c}) \\ log(m^{Cancer}_{t,s,c}) \\ log(m^{Circulatory}_{t,s,c}) \\ log(m^{Respiratory}_{t,s,c}) \\ log(m^{External}_{t,s,c}) \end{pmatrix},$$

where n is the number of the analyzed cause-specific mortality rates, here $n = 5$, t denotes the time, s the gender, and c the country.

[1] A comparability ratio serves to remove the discontinuities between the observation periods: it makes the average of the mortality rates over the last two years of a classification coincide with the average of the mortality rates over the first two years of the next classification. So, the mortality rates in every classification are divided by the comparability ratio(s) linking this classification to the previous one(s).

3. Theoretical framework

3.1. VECM and the common stochastic factors

Arnold and Sherris (2015, 2016) showed that the cause-specific mortality rates were non-stationary and so, contained stochastic trends. It was also demonstrated that at least one cointegration relation existed between the variables and for this reason, it was possible to build a Vector Error Correction Model (VECM) describing the development of the cause-specific mortality rates. VECMs represent an extension of the Vector AutoRegression (VAR) models and allow modeling the dependency between the lagged values of the differenced variables and the variables in levels through the so-called cointegration term $\alpha\beta'\mathbf{y}_{t-1}$. Supposing that there are r cointegration relations, i.e. that there exists a matrix β of rank r such that $\beta'\mathbf{y}_t$ is I(0), the corresponding VECM has the following form:

$$\Delta\mathbf{y}_t = \mathbf{c} + \mathbf{d}t + \alpha\beta'\mathbf{y}_{t-1} + \sum_{i=1}^{p-1} \xi_i \Delta\mathbf{y}_{t-i} + \epsilon_t, \quad t = 1...T \tag{2}$$

where

- $\mathbf{c}$ and $\mathbf{d}$ are $(n \times 1)$ vectors of constants;
- ξ_i is a $(n \times n)$ matrix of autoregressive coefficients for $i = 1, 2, \ldots, p-1$;
- β is a $(n \times r)$ matrix containing r vectors each representing a cointegration relation;
- α is a $(n \times r)$ loading matrix that indicates how a particular variable is impacted by the cointegration relation;
- ϵ_t is a $(n \times 1)$ vector of white noise errors.

Hamilton (1994) and Lütkepohl (2005) are the extensive references on the VECM and VAR models.

Further, as mentioned in Stock and Watson (1988), cointegrated multivariate time series comprise at least one common trend and so, can be expressed as a sum of a reduced number of common stochastic trends, plus transitory, or stationary, components. In other words, $\mathbf{y}_t$ can be explained in terms of a smaller number $(n - r)$ of I(1) variables, f_t, called common factors or common long-memory components, plus some I(0) components $\tilde{\mathbf{y}}_t$:

$$\underset{n\times 1}{\mathbf{y_t}} = \underset{n\times k}{A_1}\ \underset{k\times 1}{f_t} + \underset{n\times 1}{\tilde{\mathbf{y}}_\mathbf{t}}, \tag{3}$$

where $k = n - r$.

The objective is then to estimate f_t as a linear combination of the original variables using the methodology developed by Gonzalo and Granger (1995) which we briefly present in Appendix C. Although the f_t have fewer number of dimensions than the original data, further reduction of dimensionality may be needed to allow comparing of the common factors across countries.

3.2. Principal component analysis

One of the most popular dimension-reduction techniques is the principal component analysis (PCA) and the book by Jolliffe (2002) is an extensive reference on the subject. Simply put, the idea of the PCA is to reduce the dimensionality of a dataset and, in the same time, preserve as much as possible the variation present in the data.

Suppose that $\mathbf{y}$ is a vector of n random variables with a known covariance matrix Σ. If Σ is a positive definite, it can be decomposed as $\Sigma = \Gamma'\Lambda\Gamma$, where columns $v_1, v_2, \ldots, v_n$ of Γ are the eigenvectors corresponding to the ordered eigenvalues $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$ which form the main diagonal of the matrix Λ. It can be shown that the vector $\mathbf{x} = \Gamma\mathbf{y}$ will have the

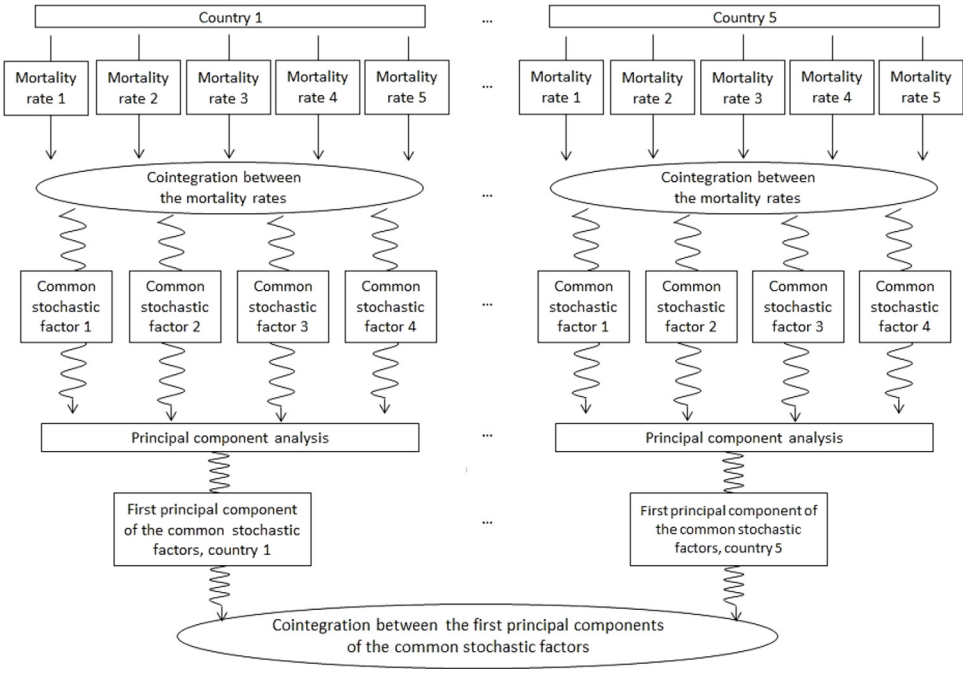

Fig. 1. Two-level cointegration analysis of the cause-specific mortality rates.

same total variance as the vector **y** and the first component of **x** given by $x_1 = v_1'\mathbf{y}$ will have the maximum variance of any linear combination $a'\mathbf{y}$ such that $||a|| = 1$ (Hogg et al., 2014). For this reason, x_1 is called the first principal component (PC) of **y** and can be used as a proxy for the information contained in **y**.

In our data, some elements of the vector **y** substantially outweigh the rest of the causes (e.g. circulatory and cancer death rates). In this case it is recommended to use the correlation matrix Σ^* of **y** when calculating the PCs (Jolliffe, 2002).

So we apply the PCA in order to express the information contained in the common stochastic factors in one principal component. Such a reduction in the dimensionality of the data will allow us to compare the first principal components for different countries and sexes using the cointegration tests. This process is schematically shown in Fig. 1.

Added to above, the PCA is a well-known method in the mortality modeling field. One of the most popular mortality models, the Lee–Carter model, basically extracts and projects a unique time trend from a matrix of mortality rates by assuming that the vector of the mortality reductions is time-invariant. Yang et al. (2010) go further and account for the variant mortality improvements at different ages.

4. Application to the cause-specific mortality rates

4.1. Estimation of the common stochastic factors

In order to estimate the common stochastic factors, one has to, first, find the VECM equations that best describe the datasets. Here, not only the number of cointegrating relations, but also the form of the deterministic part of the model play an important role. Let $\mu_t = \mathbf{c} + \mathbf{d}t$ denote the deterministic part of the model (2) and suppose that the parameter d can be decomposed in the directions of $\delta_\perp$ and δ such that $\delta\delta_\perp = 0$. Then $\mathbf{d} = \delta\rho + \delta_\perp\gamma$, where ρ and γ are the decomposition parameters. As the mortality rates are known to have a trend, we will consider the following forms of the deterministic elements (Johansen, 1995):

- NT: no trend in the VECM, but a linear trend in the levels of the variables: $\mathbf{c} \neq 0, \rho = 0, \gamma = 0$, hence $\mathbf{d} = 0$,
- TC: linear trend in the cointegration relation combined with a linear trend in the levels of the variables (i.e., no linear trend in the differenced variables): $\mathbf{c} \neq 0, \rho \neq 0, \gamma = 0$, hence $\mathbf{d} = \delta\rho$,
- QT: linear trend in the differenced variables, thus the quadratic trend in the levels of the variables (i.e., the VAR model) : $\mathbf{c} \neq 0, \rho \neq 0, \gamma \neq 0$, hence $\mathbf{d} = \delta\rho + \delta_\perp\gamma$.

Table 1
Vector error correction models chosen for the analysis.

Country	Males	Females
US	VAR(2), QT, 1 CR	VAR(2), QT, 1 CR
JP	VAR(2), TC, 2 CR	VAR(2), TC, 2 CR
FR	VAR(2), NT, 1 CR	VAR(2), QT, 1 CR
E&W	VAR(2), QT, 1 CR	VAR(2), QT, 1 CR
AU	VAR(2), QT, 1 CR	VAR(2), NT, 1 CR

Note: QT = quadratic trend in the VAR; TC = linear trend in the cointegration relation; NT = no trend; CR = cointegration relation.

We will use the same VECMs as in Arnold and Glushko (2021) that we reproduce here for the sake of completeness (see Table 1).

Once the VECM coefficients are calculated for every dataset, we estimate the common factors $f_t = \alpha_\perp\ \mathbf{y_t}$ as described in the previous section. The maximum likelihood estimates of $\hat{\alpha}_\perp = (\hat{m}_{r+1}, \ldots, \hat{m}_n)$, where r is a number of cointegration relations, that is 1 or 2 for the datasets used in the study, suggest that the number of dimensions of the common factor component f_t will be 4 or 3 respectively. Hence, by estimating common stochastic factors we reduced the number of dimensions in our system, but comparing common factors across countries remains complicated. Fig. 2 shows the common factors estimated for the dataset of US males, common factors for the rest of the datasets are shown on Figs. B.1–B.9 of Appendix B.

At present, we would like to compare the common stochastic factors across five countries, but as we have 3 to 4 factors for each country, we cannot apply the cointegration analysis. For this, we need to further reduce the number of dimensions in order to be able to compare the common trends across the different datasets, and for this, we use the principal component analysis.

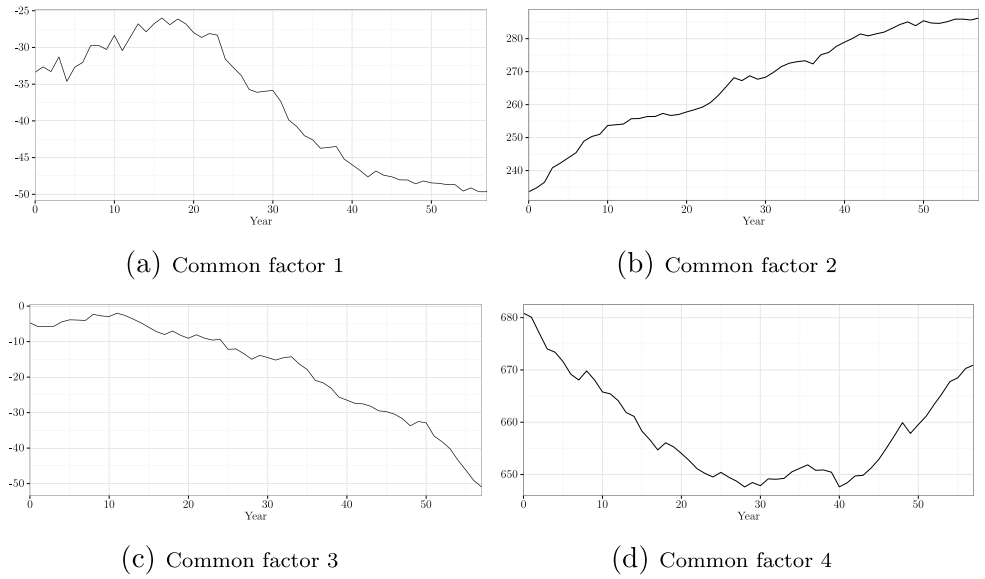

Fig. 2. Common stochastic factors f_t, US males.

Table 2
Proportion of the variance explained by the first principal component.

	US	JP	FR	EW	AU
Males	72%	64%	79%	71%	79%
Females	79%	91%	76%	74%	90%

Table 3
Test for the number of cointegration relations, male datasets.

Trace statistics		Critical values (Case NT)			
r	Males	10%	5%	2.5%	1%
4	3.50	2.69	3.76	4.95	6.65
3	11.75	13.33	15.41	17.52	20.04
2	25.78	26.79	29.68	32.56	35.65
1	44.80	43.95	47.21	50.35	54.56
0	83.14	64.84	68.52	71.80	76.07
Max eigenvalue statistics		**Critical values (Case NT)**			
r	Males	10%	5%	2.5%	1%
4	3.50	2.69	3.76	4.95	6.65
3	8.25	12.07	14.07	16.05	18.63
2	14.03	18.60	20.97	23.09	25.52
1	19.02	24.73	27.07	28.98	32.24
0	38.34	30.90	33.46	35.71	38.77

A null hypothesis is accepted at α% significance level when the statistic is lower than the corresponding critical value. Hence, the hypothesis of r equal to 1 is accepted by both tests at 5%, 2.5% and 1%.

The principal components for the US males dataset are shown in Fig. 3 (see Figs. B.10–B.18 of Appendix B for the rest of the countries and sexes). As expected, the first PC has the maximum variance, whereas the rest of the components represent fluctuations around the zero line.

For US males, the first principal component explains approximately 72% of the total variance. Table 2 shows proportions for the remaining datasets: for all datasets except Japanese males the first principal component accounts for at least 70% of the total variance. For the sake of comparability, we will keep and compare the first PCs in every country and for every sex.

When comparing the charts of the principal components, it is interesting to notice that the forms of the first components on each chart have a high degree of resemblance. The similarity becomes even more striking once the first PCs from every dataset are put on the same chart (Figs. 4 and 5). To improve comparability, we multiplied some of the PCs by -1 as we know that such operations have no impact on the orthogonality or the variance accounted for by a given principal component.

Although the cause-specific mortality rates showed rather different development profiles depending on the country, we see that the patterns of the common stochastic factors, condensed using the principal component analysis, share a lot of similarities across the datasets. The resemblance is even more pronounced for the female datasets. So, we would like to measure the closeness between the first principal components of the common stochastic factors in a formal way, using the tools of the cointegration analysis. If some non-stationary variables are found to be cointegrated, then there exists their linear combination that is stationary even if each variable is not. Also, cointegrated variables move together over the long term and are subject to the influence of the same common trends. If the first PCs from every dataset are cointegrated, this will mean that they are linked to each other in their long-term development.

4.2. Cointegration between the first principal components of the common stochastic factors

In order to test if the first principal components of the common stochastic factors are cointegrated, we apply the Johansen maximum likelihood procedure (Johansen, 1988). This is possible only for the time series with the same number of observations. For this reason, we are obliged to cut the first PCs at the length of the shortest among them.

Under the null hypothesis H_0 of the *trace test* there are exactly r cointegrating relations among the data vector, in our case five first principal components from five countries, against the alternative hypothesis H_A that there are $n = 5$ cointegration relations. From the upper part of Table 3 we see that H_0 is rejected for the $r = 0$ and is accepted for $r = 1$ at 5%, 2.5% and 1% significance level. So according to the trace test there is 1 cointegration relation for the male dataset. Under the null hypothesis H_0 of the *maximum eigenvalue test* there are exactly r cointegrating relations among the data vector against the alternative hypothesis H_A that there are $r + 1$ cointegration relations. From the lower part of Table 3 we see that H_0 is rejected for the $r = 0$ and

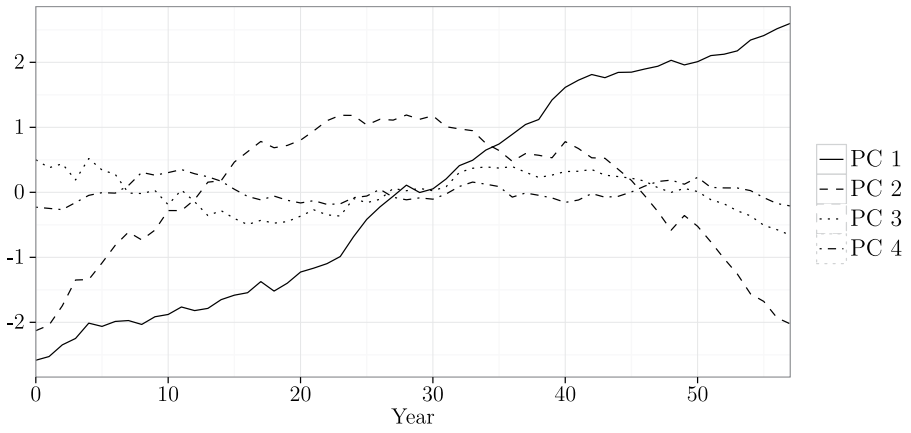

Fig. 3. PCA applied to common factors, US males.

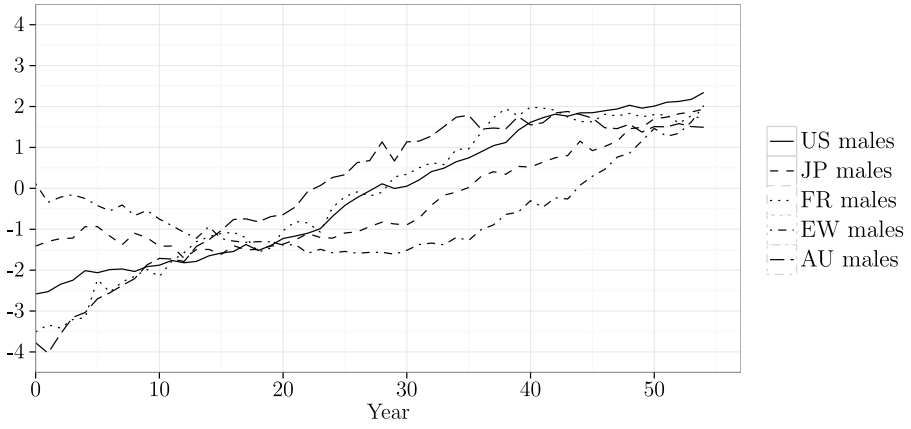

Fig. 4. First PCs for each country, males.

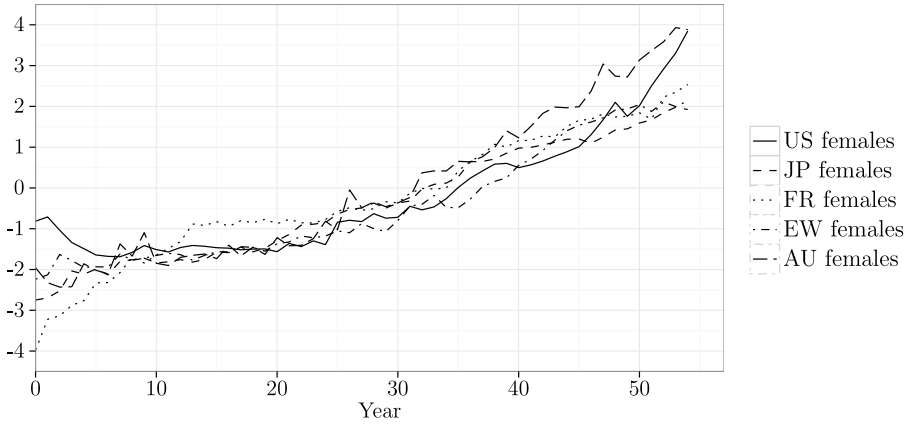

Fig. 5. First PCs for each country, females.

is accepted for $r = 1$ at all shown significance levels. We can conclude that according to both tests there is 1 cointegration relation for the male dataset. Thus, although the common factors evolve stochastically, their development is driven by one long-run equilibrium relationship that remains stationary over the years.

Similar observations hold for the female datasets with the number of cointegrating relations being equal to 2 in the NT case and to 1 in the QT case (numerical results are shown in Appendix B, Table B.1 and Table B.2).

For both male and female datasets taken separately, Johansen tests for the form of the deterministic term indicate that either no trend (case NT) or a quadratic trend (case QT) should be included in the process. For male datasets, the results of cointegration test with a quadratic trend are inconclusive, and for this reason, we

Table 4
Tests on residuals of the fitted VECM.

Type of the test	Name of the test	p value		
		Males NT $r = 1$	Females NT $r = 2$	Females QT $r = 1$
Autocorrelation	Portmanteau (15 lags)	0.348	0.970	0.991
	Portmanteau (25 lags)	0.162	0.975	0.982
	Portmanteau (35 lags)	0.159	0.985	0.991
Normality	Skewness	0.335	0.653	0.180
	Kurtosis	0.311	0.967	0.603
	Both	0.309	0.935	0.339

present the results for the NT case only. For female datasets, both forms of the deterministic term provide good model fit. The null hypotheses of no autocorrelation and normality of the residuals are not rejected for the male, as well as the female dataset (Table 4).

We also tested the male and female datasets together for the cointegration between the first principal components of the common stochastic factors. As the shape of the first PCs is quite different between the male and female datasets, it is not surprising that we do not find any cointegration relation when we test the first principal components of the common stochastic factors for both sexes simultaneously.

As for the first principal components of the common stochastic factors tested separately for each sex, they are indeed cointegrated. This means that first, the cause-specific mortality rates, being cointegrated themselves, share some stochastic trends that are common to all causes. Then, once the corresponding common stochastic factors are explicitly extracted from the cause-specific mortality rates and condensed, they, in turn, also share some common stochastic trends, but on the next level, i.e. across different countries.

5. Discussion and conclusion

Although cause-specific mortality rates show different development patterns across countries, sexes and historical periods, as we deal with the death rates of the human species, it is reasonable to expect that similarities and common features also exist between these patterns. The steady decrease of the mortality rates that has been observed in many parts of the world for more than a century is also due to factors and effects that are universally present, albeit to a different extent. Among such factors one can cite medical advances, changes in lifestyles and nutrition, epidemics, and aging. Cointegration analysis in conjunction with the identification of the common stochastic factors as proposed by Gonzalo and Granger (1995) is then a practical tool that can efficiently help to elicit possible common long-term regularities and trends.

Among the factors influencing the development of mortality rates, biological aging has probably the most universal character as it increases vulnerability to all common causes of death (Olshansky et al., 2002; Hayflick, 2004). Although there is no single generally accepted definition of aging, in simple terms it is usually defined as a progressive loss of normal body functions that leads to death (Holliday, 2004). In biological systems, aging is believed to be a stochastic process that is caused by the increasing loss of molecular fidelity that, with time, becomes superior to the repair capacity of the organisms. Also, age changes represent the greatest risk factor for age-associated diseases (Hayflick, 2004).

In spite of its paramount importance to virtually every aspect of human life, the process of aging is not yet fully understood and many different theories of aging exist, each having its merits and weaknesses (for a review thereof see Holliday (2004) and Jin (2010) including the references). Moreover, it is still not clear if aging can be measured since a reliable biomarker of aging is yet to be found (Butler et al., 2004). Should such measure be discovered, we could then study the development of mortality rates in function of advancement in aging. Instead, we presently find ourselves in a situation when a set of mortality rates evolving in time *must* reflect the effect of the aging processes, but it is not clear if and how this effect can be made explicit on the basis of the observed mortality rates only.

The intuition behind the application of the cointegration analysis to the death rates is to model simultaneously the development of several main cause-specific mortality rates and repeating the analysis for a number of populations. Should any patterns or trends be revealed that are common to all or the majority of the datasets included in the study, this could point to some fundamental processes that are proper to the human species. In Arnold and Sherris (2016) the authors found that only the endogenous mortality rates participated in the cointegration relations, i.e. in the long-term equilibrium states between the causes. It was assumed that this could indicate the link between the common stochastic trends shared by every cause-specific mortality rate and the processes of biological aging.

In this study, we explicitly measured the common stochastic factors as well as proposed an approach allowing us to make comparisons across countries. For this, we first, extracted the common stochastic factors from the sets of the cause-specific mortality rates in every country and for each sex. These are the factors that impact every cause-specific mortality rate in a particular country. Then, using the principal component analysis we condensed these factors and used the first principal components in the subsequent cross-country analysis. We have found that there exists at least one cointegration relation between the first principal components of the common stochastic factors from different countries, which means that they are also subject to some universal stochastic trends that deploy their impact in all countries. As a consequence, these universal stochastic trends might reflect some intrinsic processes that occur in every country. Our results, combined with those of Arnold and Sherris (2016) tend to indicate that these universal stochastic trends are describing some features of the processes of biological aging, although at this point, we cannot state with certainty to what exactly these trends correspond. As we now know that they exist and can be made explicit from the data, further research is needed in order to identify the mechanisms, likely biological, that are behind the observed behavior of the cause-specific mortality rates. Because of its universal and predominant character, aging is one of the possibilities that should not be omitted.

We believe that our results will bring forward the discussion on how to measure the biological processes of aging and the related research in the fields of demography, economics, biology or epidemiology. In addition, our study shows that as the same stochastic trends are present in all datasets, similar assumptions for the intra-cause dependence can be used across different countries. On the other hand, we saw that the first principal components of the common stochastic factors for males and females when put together were not cointegrated. This could indicate that there is an important divergence in the biology of aging for men and women and for this reason, different assumptions should be used for each sex when modeling the development of the cause-specific mortality rates. Our next steps will consist in using the results of the cointegration analysis to improve the cause-specific mortality forecasts.

Acknowledgments

Séverine Arnold and Viktoriya Glushko gratefully acknowledge the financial support received from the Swiss National Science Foundation for the project "Cause-Specific Mortality Interactions" (project number 100018_162898). The authors also would

Table A.1
Age-standardized central death rates for selected years, x10^3.

		Females			Males		
		1960	1980	2000	1960	1980	2000
US	I&P	0.0852	0.0630	0.1014	0.1888	0.1215	0.1686
	Cancer	1.6269	1.5857	1.5786	2.1394	2.4409	2.2483
	Circulatory	4.3148	2.9428	1.9744	6.8425	5.2315	3.2263
	Respiratory	0.4595	0.3292	0.5124	0.8670	0.8205	0.8512
	External	0.4516	0.3804	0.3058	1.1217	1.0368	0.7731
JP	I&P	0.4350	0.0842	0.0629	0.8160	0.2099	0.1426
	Cancer	1.4299	1.3156	1.1361	2.0147	2.1748	2.1092
	Circulatory	1.6850	1.3904	0.6263	2.2670	1.8094	0.8788
	Respiratory	0.9022	0.3804	0.3551	1.3673	0.6830	0.7941
	External	0.4644	0.3230	0.2476	1.0982	0.7577	0.7161
FR	I&P	0.1759	0.1085	0.0833	0.4487	0.2171	0.1502
	Cancer	1.6852	1.5080	1.3868	2.5972	3.0608	2.8610
	Circulatory	2.1922	1.6042	0.9415	3.3823	2.7013	1.6697
	Respiratory	0.8405	0.3023	0.3385	1.3661	0.6451	0.5930
	External	0.5020	0.5267	0.3623	1.1683	1.1520	0.8470
EW	I&P	0.0721	0.0387	0.0557	0.1764	0.0636	0.0708
	Cancer	1.7114	1.7764	1.5838	2.6498	2.8131	2.2807
	Circulatory	3.7128	2.7551	1.4769	5.7063	4.9203	2.5757
	Respiratory	0.8536	0.9398	0.8048	1.9318	1.8931	1.2785
	External	0.3823	0.3044	0.1780	0.6787	0.5561	0.4205
AU	I&P	0.0695	0.0296	0.0419	0.1518	0.0535	0.0843
	Cancer	1.4781	1.4868	1.4368	2.0721	2.3765	2.1570
	Circulatory	4.0123	3.0032	1.5502	6.5525	5.1796	2.4928
	Respiratory	0.5108	0.4362	0.6109	1.1511	1.1897	1.0071
	External	0.4747	0.4236	0.2625	1.0533	0.9189	0.6118

like to thank an anonymous referee for the useful comments which helped to improve the article. All error are their sole responsibility.

Appendix A

Age-standardization of the cause-specific mortality rates

To calculate the age-standardized death rates, we first, calculate the simple mortality rates for each country as the number of deaths by age, sex and cause divided by the mid-year population by age and sex. Next, we assume that the population age structure is constant over the whole observation period and is equal to the age structure of the US males population in 2007 (see Table A.1):

$$m_{x,t,d,s,c} = d_{x,t,d,s,c}/l_{x,t,s,c}$$

$$d^*_{t,d,s,c} = \sum_x m_{x,t,d,s,c} \times l_{x,2007,males,USA}$$

$$m^*_{t,d,s,c} = d^*_{t,d,s,c}/l_{2007,males,USA}$$

where (see Fig. A.1).

$d_{x,t,d,s,c}$ = number of deaths at age x, in year t, for cause of death d, gender s and country c;

$l_{x,t,s,c}$ = mid-year population at age x, in year t, gender s and country c;

$m_{x,t,d,s,c}$ = central death rate at age x, in year t, for cause of death d, gender s and country c.

Appendix B

See Figs. B.1–B.18, Tables B.1 and B.2.

Table B.1
Test for the number of cointegration relations, female datasets, case NT.

Trace statistics		Critical values (NT)			
r	Females	10%	5%	2.5%	1%
4	0.79	2.69	3.76	4.95	6.65
3	6.16	13.33	15.41	17.52	20.04
2	18.18	26.79	29.68	32.56	35.65
1	52.39	43.95	47.21	50.35	54.56
0	101.41	64.84	68.52	71.80	76.07
Max eigenvalue statistics		Critical values (NT)			
r	Females	10%	5%	2.5%	1%
4	0.79	2.69	3.76	4.95	6.65
3	5.37	12.07	14.07	16.05	18.63
2	12.02	18.60	20.97	23.09	25.52
1	34.20	24.73	27.07	28.98	32.24
0	49.02	30.90	33.46	35.71	38.77

A null hypothesis is accepted at α% significance level when the statistic is lower than the corresponding critical value. Hence, the hypothesis of r equal to 2 is accepted at all significance levels by both tests.

Table B.2
Test for the number of cointegration relations, female datasets, case QT.

Trace statistics		Critical values (QT)			
r	Females	10%	5%	2.5%	1%
4	0.05	2.57	3.74	4.85	6.40
3	5.37	16.06	18.17	20.13	23.46
2	15.75	31.42	34.55	36.94	40.49
1	35.22	50.74	54.64	57.79	61.24
0	83.85	73.40	77.74	80.74	85.78
Max eigenvalue statistics		Critical values (QT)			
r	Females	10%	5%	2.5%	1%
4	0.05	2.57	3.74	4.85	6.40
3	5.32	14.84	16.87	18.57	21.47
2	10.38	21.53	23.78	26.07	28.83
1	19.47	27.76	30.33	32.56	35.68
0	48.64	33.74	36.41	38.68	41.58

A null hypothesis is accepted at α% significance level when the statistic is lower than the corresponding critical value. Hence, the hypothesis of r equal to 1 is accepted at all significance levels by both tests.

Appendix C

Estimation of the common stochastic factors

The methodology developed by Gonzalo and Granger (1995) allows to estimate f_t as defined in (3) by imposing that

1. f_t be linear combinations of the original variables, in our case the cause-specific mortality rates:

$$\underset{k\times 1}{f_t} = \underset{k\times n}{B_1}\ \underset{n\times 1}{\mathbf{y_t}}, \tag{C.1}$$

2. the remaining stationary part $\tilde{\mathbf{y}}_\mathbf{t}$ does not have any permanent effect on $\mathbf{y_t}$.

Substituting (C.1) in (3), we obtain that $\tilde{\mathbf{y}}_\mathbf{t} = (I - A_1B_1)\mathbf{y_t}$. In other words, the stationary component $\tilde{\mathbf{y}}_\mathbf{t}$ is also a linear combination of the non-stationary variables $\mathbf{y_t}$ which is only possible for $\tilde{\mathbf{y}}_\mathbf{t} = A_2\beta'\mathbf{y_t} = A_2z_t$, where $z_t = \beta'\mathbf{y_t}$ is the cointegration relation. The authors show that the only linear combination of $\mathbf{y_t}$ such that $\tilde{\mathbf{y}}_\mathbf{t}$ has no long-run effect on $\mathbf{y_t}$ is

$$f_t = \underset{k\times n}{\alpha_\perp}\ \underset{n\times 1}{\mathbf{y_t}}, \tag{C.2}$$

where $\alpha'_\perp\alpha = 0$ and $k = n - r$.

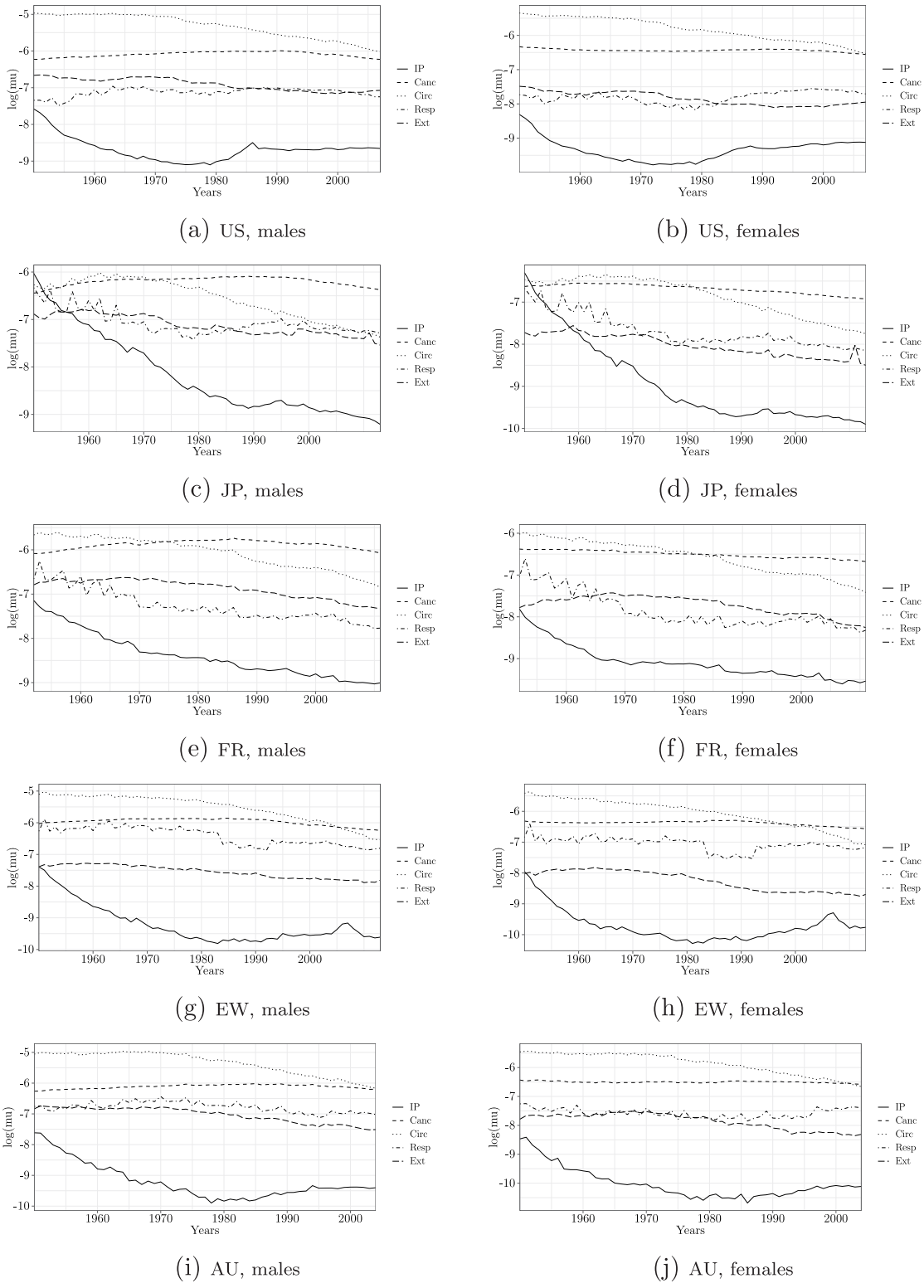

Fig. A.1. Age-standardized central log-death rates by cause.

The condition imposed in (C.1) not only helps to identify f_t, but also makes them observable by linking f_t to the original variables. Both conditions make f_t "a good candidate to summarize the long-run behavior of the original variables" (Gonzalo and Granger, 1995). The authors also show that these conditions allow identifying f_t up to a non-singular matrix multiplication to the left. The resulting factor model is:

$$\mathbf{y_t} = A_1 f_t + A_2 z_t, \tag{C.3}$$

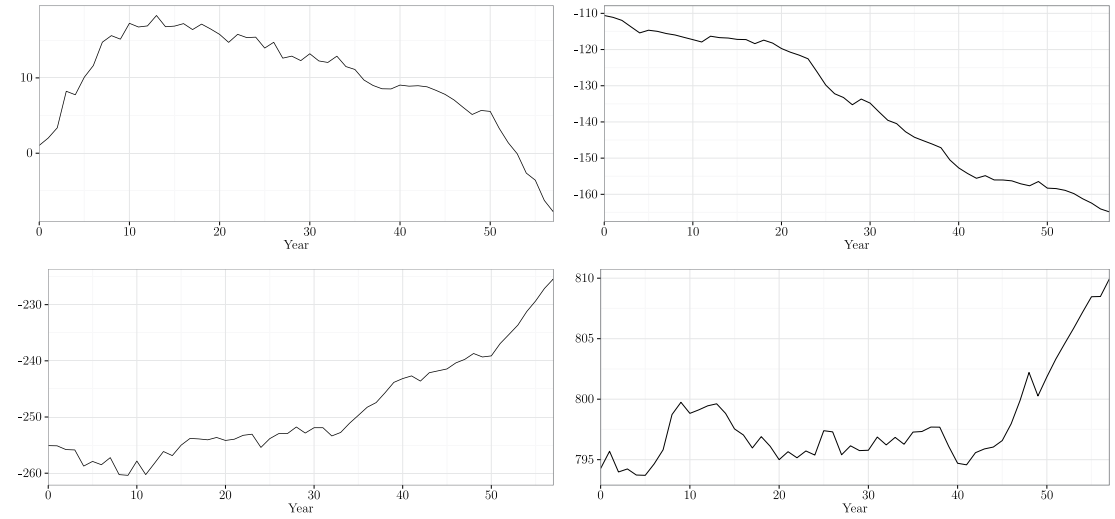

Fig. B.1. Common stochastic factors, US females.

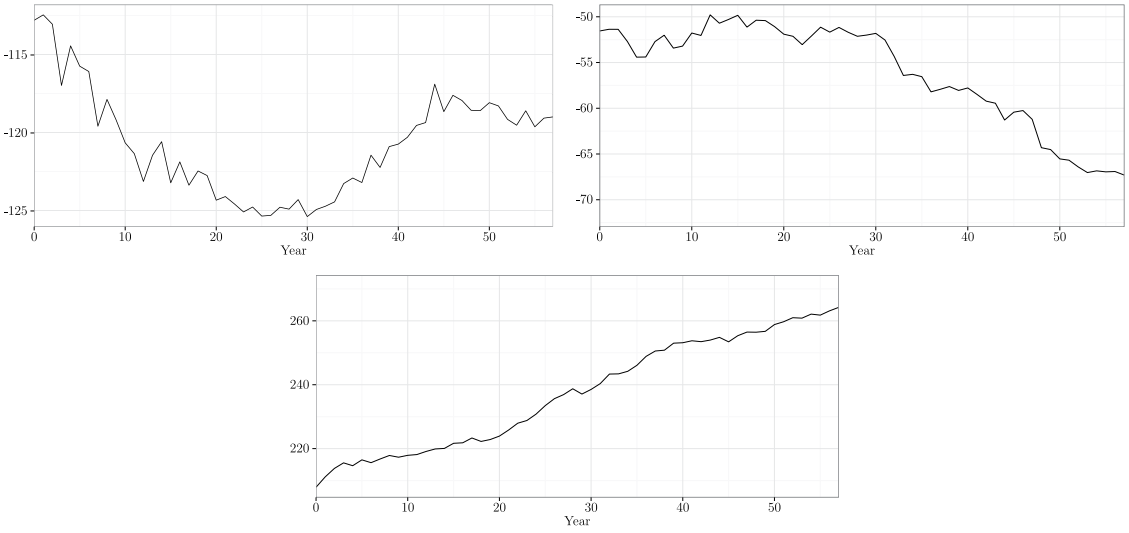

Fig. B.2. Common stochastic factors, JP males.

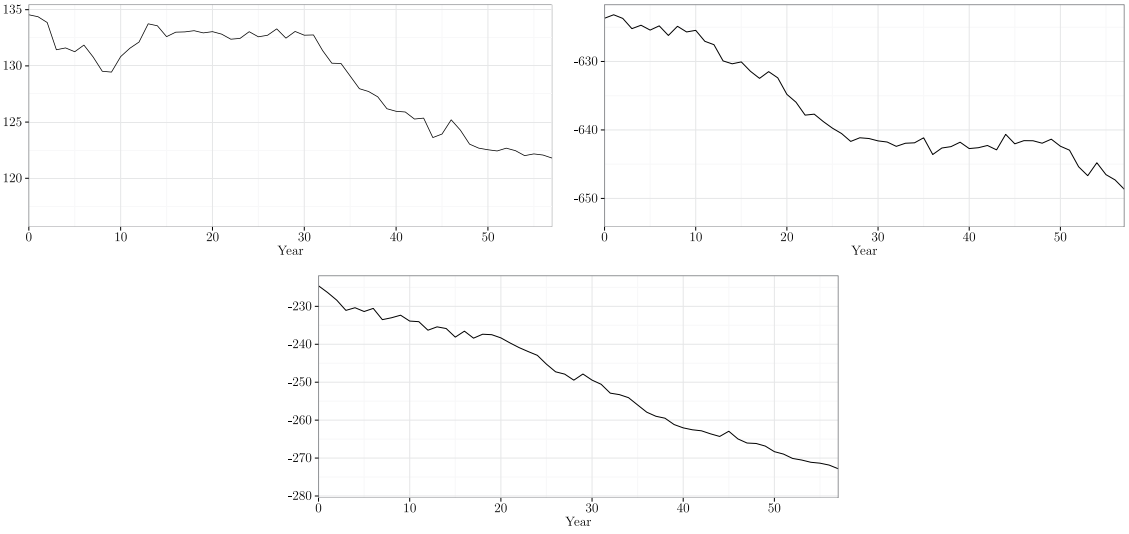

Fig. B.3. Common stochastic factors, JP females.

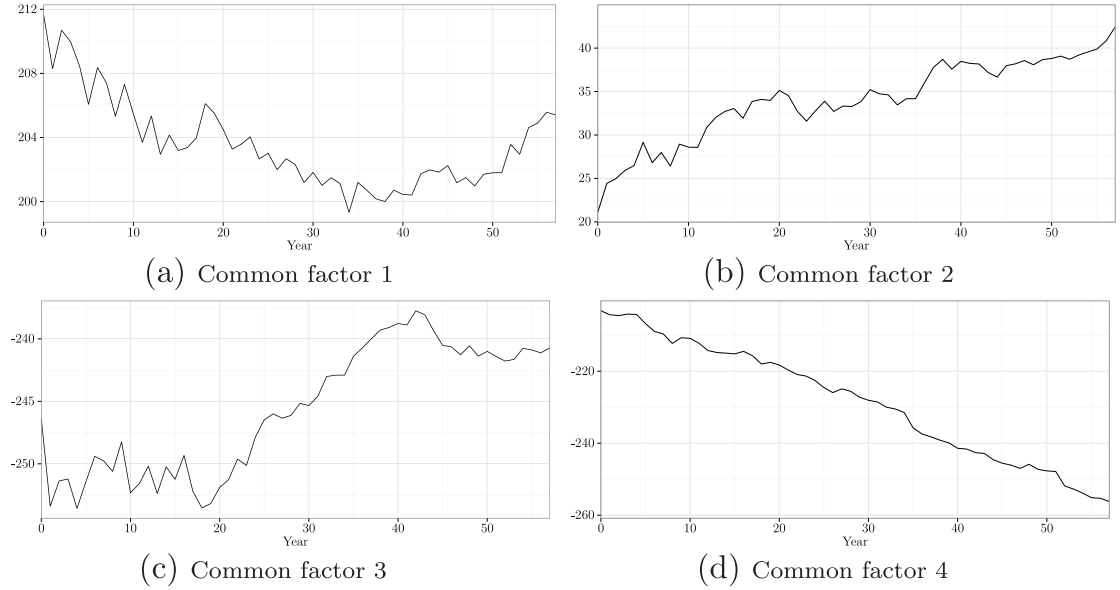

Fig. B.4. Common stochastic factors, FR males.

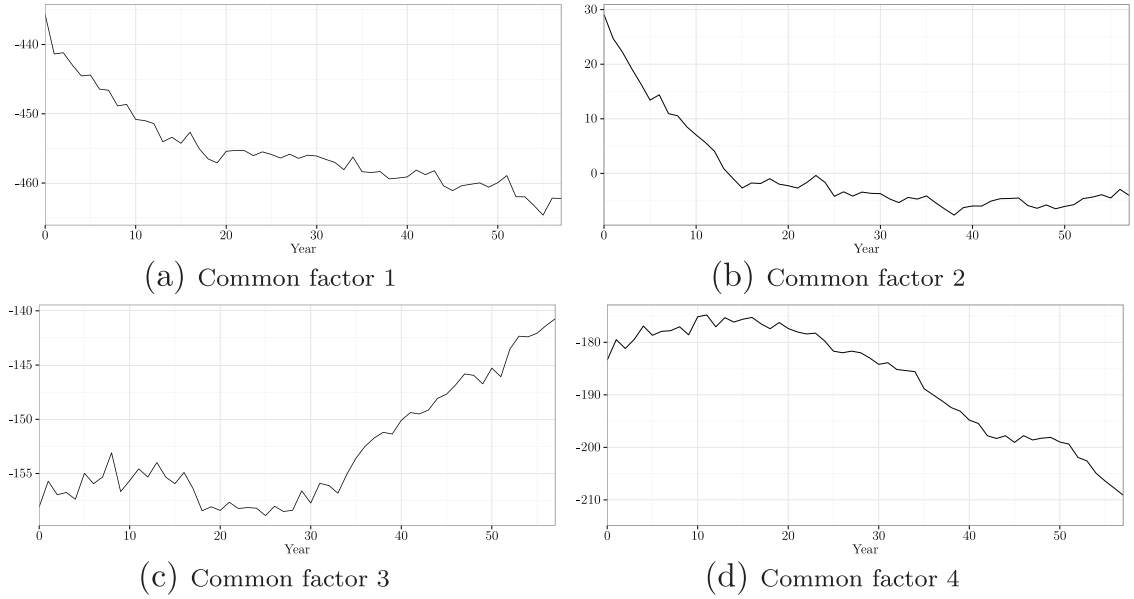

Fig. B.5. Common stochastic factors, FR females.

where $f_t = \alpha_\perp \mathbf{y_t}$ and $z_t = \beta' \mathbf{y_t}$, and satisfies the following properties:

- The common factors f_t are not cointegrated.
- $Cov(\Delta f^*_{it}, z^*_{j,t-s}) = 0$ $(i = 1, \ldots, k; j = 1, \ldots, n-k; s \geq 0)$, where $\Delta f^*_{it} = \Delta f_{it} - E(\Delta f_{it} \mid \text{lags}(\Delta \mathbf{y_{t-1}}))$ and $\Delta z^*_{it} = \Delta z_{it} - E(\Delta z_{it} \mid \text{lags}(\Delta \mathbf{y_{t-1}}))$

The second property is another way of saying that z_t does not cause f_t in the long run. It also follows that any alternative definition of f_t will vary only by I(0) components and therefore will be cointegrated.

To solve for the coefficients of (2), following Johansen (1988) we concentrate the model by regressing $\Delta \mathbf{y}_t$ and $\mathbf{y}_{t-1}$ on $(\Delta \mathbf{y}_{t-1}, \ldots, \Delta \mathbf{y}_{t-p+1})$ which gives the residuals R_{0t} and R_{1t} respectively, as well as the residual product matrices S_{ij}

$$S_{ij} = T^{-1} \sum_{j=1}^{T} R_{0t} R_{1t}, \ i, j = 0, 1. \tag{C.4}$$

The concentrated model is then

$$R_{0t} = \alpha \beta' R_{1t} + \epsilon_t, \tag{C.5}$$

and β is estimated using the reduced-rank regression from the following eigenvalues problem

$$\mid \lambda S_{11} - S_{10} S_{00}^{-1} S_{01} \mid = 0. \tag{C.6}$$

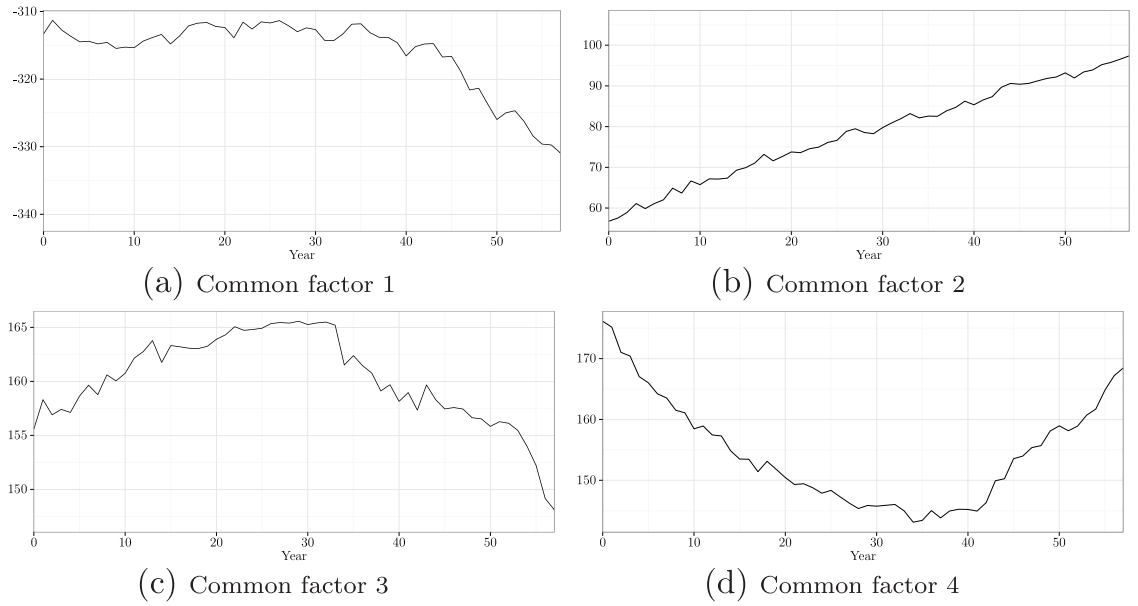

Fig. B.6. Common stochastic factors, EW males.

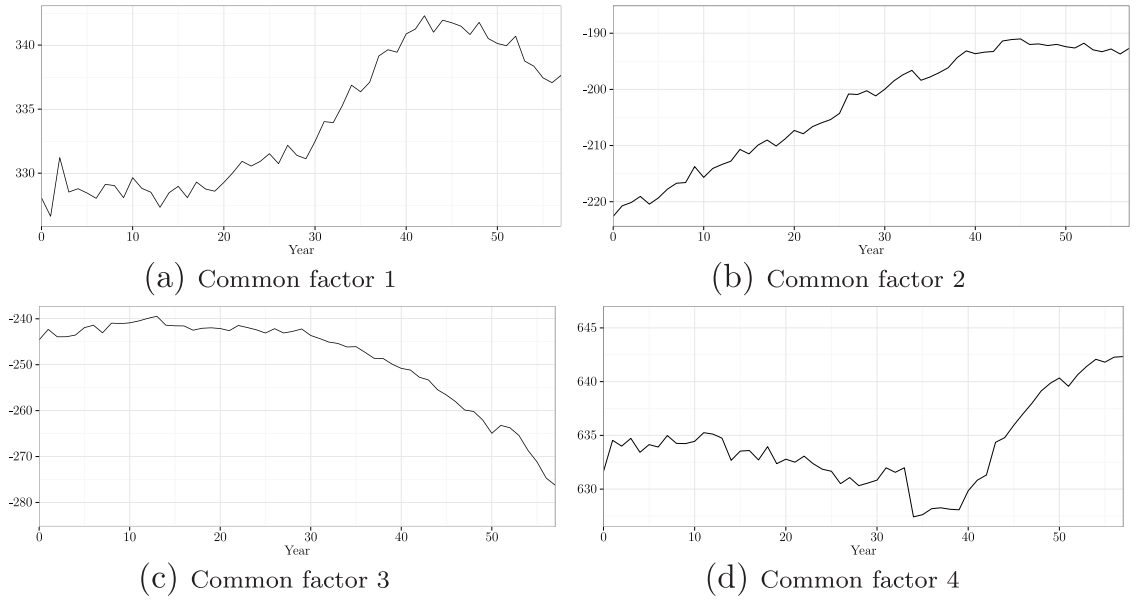

Fig. B.7. Common stochastic factors, EW females.

After ordering the eigenvalues $\hat{\lambda}_1 > \hat{\lambda}_2 > \cdots > \hat{\lambda}_n$ and corresponding eigenvectors $\hat{V} = (\hat{v}_1, \hat{v}_2, \ldots, \hat{v}_n)$, the maximum likelihood estimates of the cointegration term of the VECM are obtained as $\hat{\beta} = (\hat{v}_1, \hat{v}_2, \ldots, \hat{v}_r)$ and $\hat{\alpha} = S_{01}\hat{\beta}$.

We proceed in a similar way to estimate $\alpha_\perp$ by solving the equation

$$| \lambda S_{00} - S_{01}S_{11}^{-1}S_{10} | = 0, \tag{C.7}$$

which gives the eigenvalues $\hat{\lambda}_1 > \hat{\lambda}_2 > \cdots > \hat{\lambda}_n$ and corresponding eigenvectors $\hat{M} = (\hat{m}_1, \hat{m}_2, \ldots, \hat{m}_n)$, normalized such that $\hat{M}'S_{00}\hat{M} = I$. The $\alpha_\perp$ that defines f_t is then

$$\alpha_\perp = (\hat{m}_{r+1}, \ldots, \hat{m}_n). \tag{C.8}$$

We can see that the set of the common factors f_t has indeed a lower number of dimensions than the initial data, but if r is equal to 1 or 2, further reduction of dimensionality may be needed.

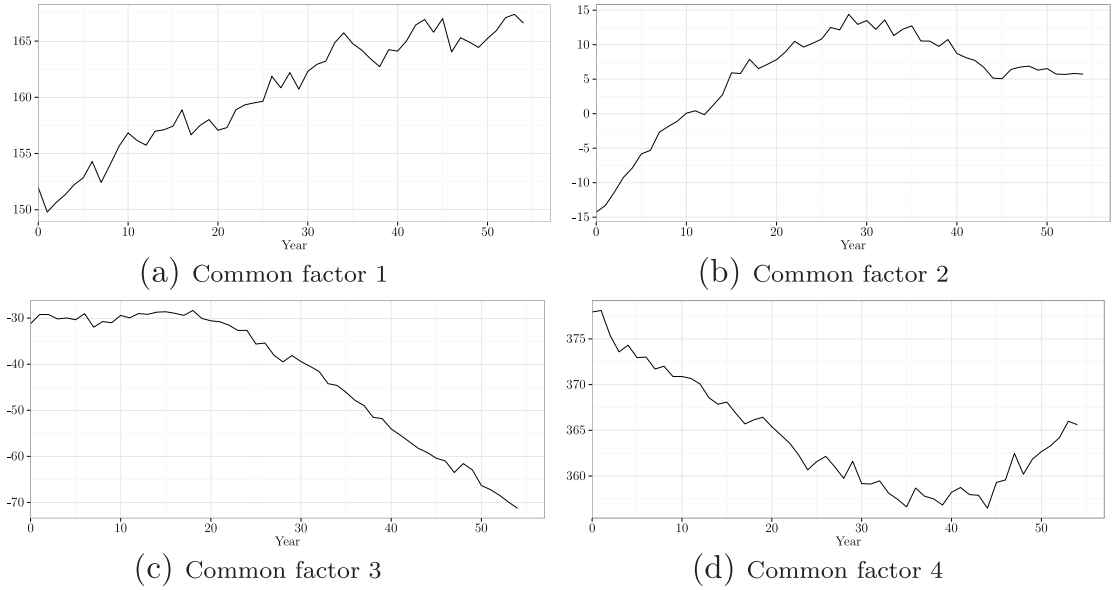

Fig. B.8. Common stochastic factors, AU males.

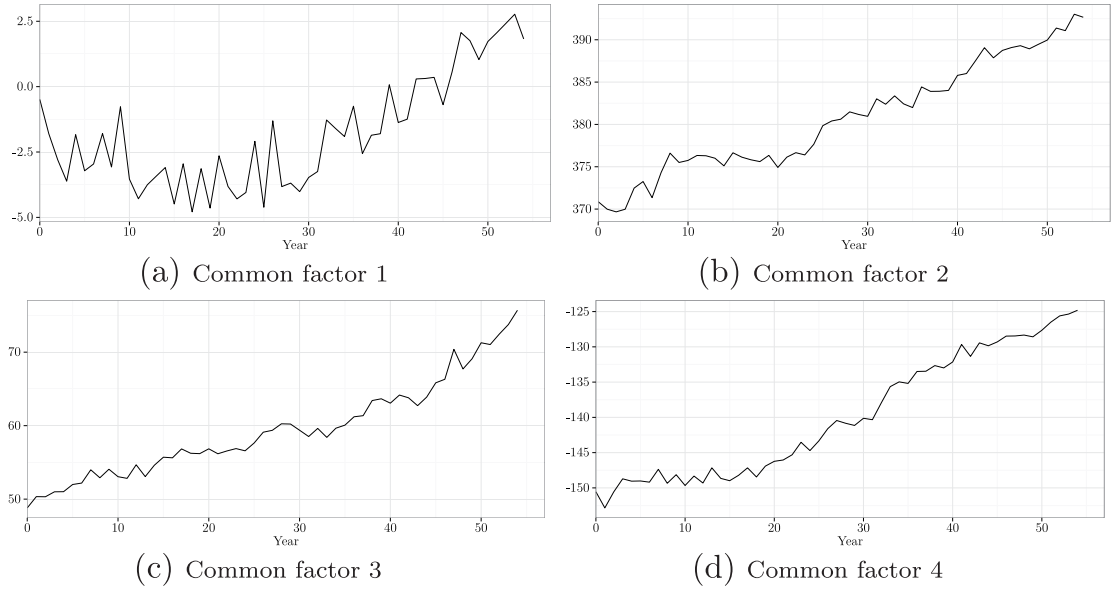

Fig. B.9. Common stochastic factors, AU females.

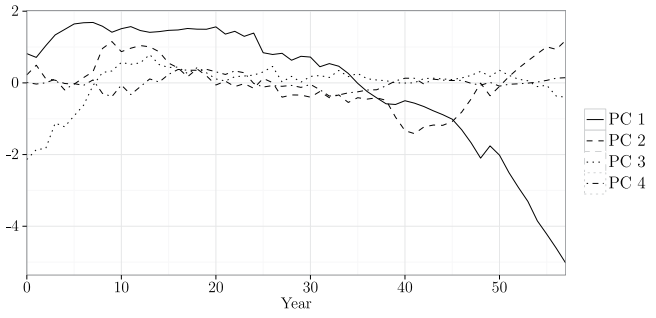

Fig. B.10. PCA applied to common factors, US females.

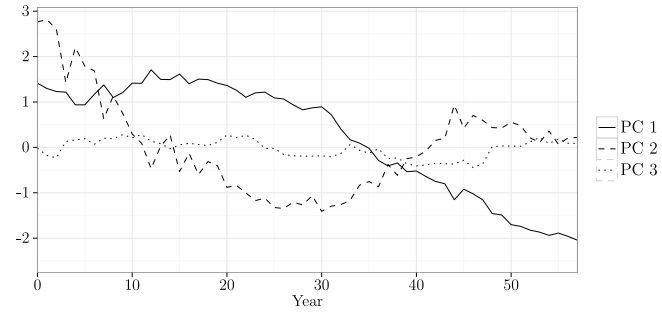

Fig. B.11. PCA applied to common factors, JP males.

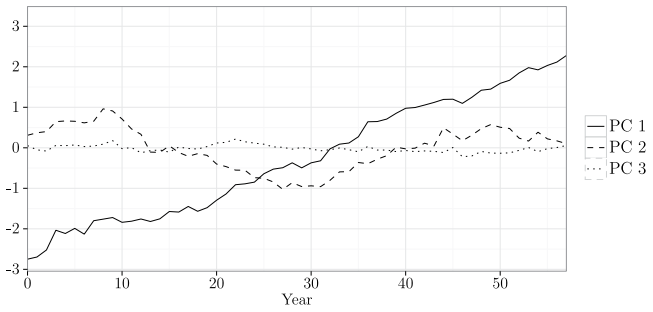

Fig. B.12. PCA applied to common factors, JP females.

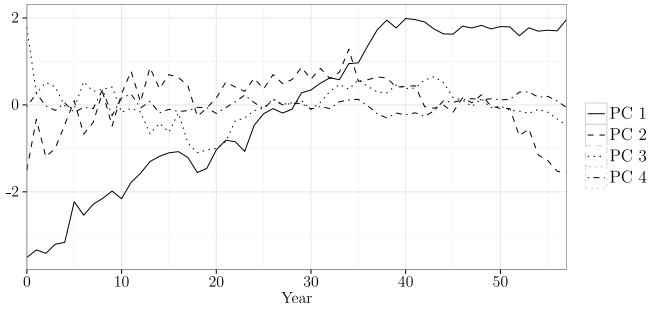

Fig. B.13. PCA applied to common factors, FR males.

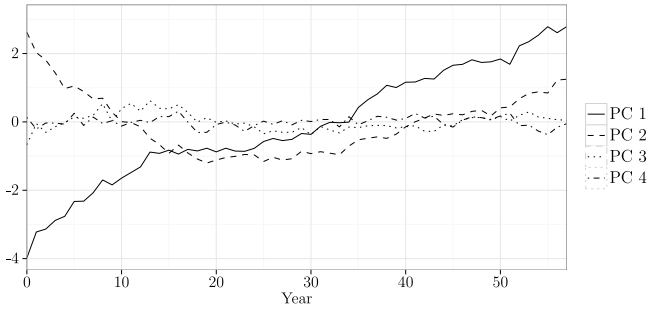

Fig. B.14. PCA applied to common factors, FR females.

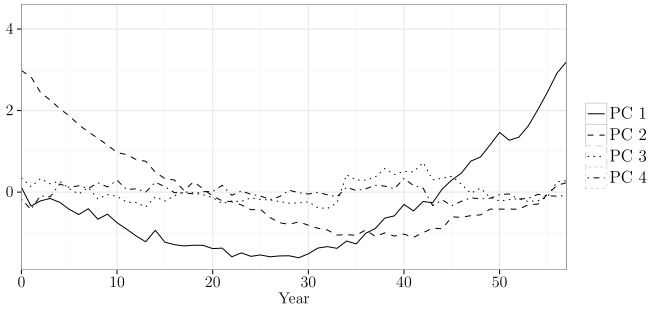

Fig. B.15. PCA applied to common factors, EW males.

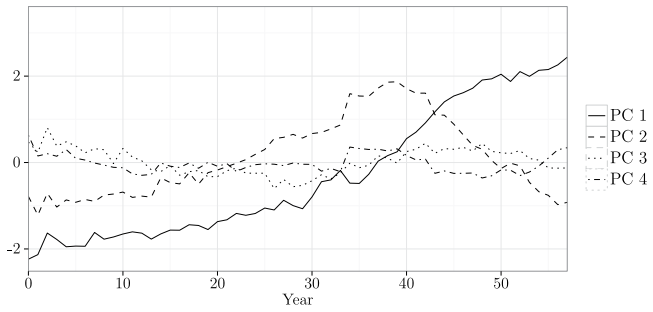

Fig. B.16. PCA applied to common factors, EW females.

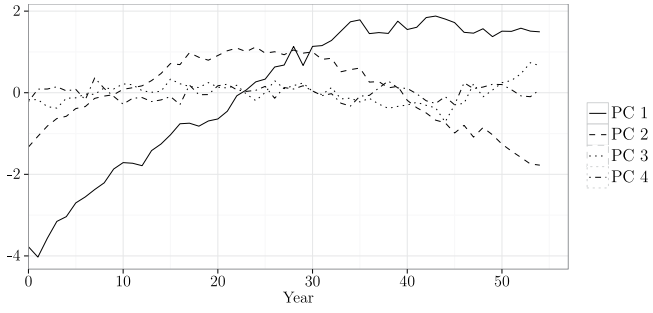

Fig. B.17. PCA applied to common factors, AU males.

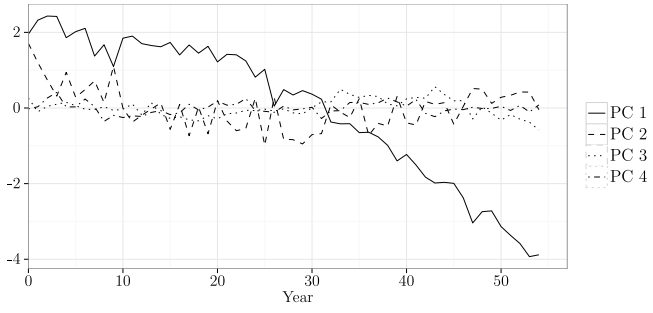

Fig. B.18. PCA applied to common factors, AU females.

References

Arnold, S., Glushko, V., 2021. Short- and long-term dynamics of cause-specific mortality rates using cointegration analysis. N. Am. Actuar. J. 0 (0), 1–23.
Arnold, S., Sherris, M., 2013. Forecasting mortality trends allowing for cause-of-death mortality dependence. N. Am. Actuar. J. 17 (4), 273–282.
Arnold, S., Sherris, M., 2015. Causes-of-death mortality: What do we know on their dependence?. N. Am. Actuar. J. 19 (2), 116–128.
Arnold, S., Sherris, M., 2016. International cause-specific mortality rates: New insights from a cointegration analysis. Astin Bull. 46 (1), 9–38.
Booth, H., Tickle, L., 2008. Mortality modelling and forecasting: A review of methods. Ann. Actuar. Sci. 3, 3–43.
Butler, R.N., Sprott, R., Warner, H., Bland, J., Feuers, R., Forster, M., Fillit, H., Harman, S.M., Hewitt, M., Hyman, M., Johnson, K., Kligman, E., McClearn, G., Nelson, J., Richardson, A., Sonntag, W., Weindruch, R., Wolf, N., 2004. Biomarkers of aging: from primitive organisms to humans. J. Gerontol.: Biol. Sci. 59A (6), 560–567.
Cairns, A., 2013. Modeling and management of longevity risk. In: Hammond, P.B., Maurer, R., Mitchell, O.S. (Eds.), Recreating Sustainable Retirement: Resilience, Solvency, and Tail Risk. Oxford University Press, Oxford, pp. 71–88.
Carnes, B.A., Holden, L.R., Olshansky, S.J., Witten, T.M., Siegel, J.S., 2006. Mortality partitions and their relevance to research on senescence. Biogerontology 7, 183–198.
Darkiewicz, G., Hoedemakers, T., 2004. How the Co-Integration Analysis Can Help in Mortality Forecasting. DTEW Research Report OR 0406, K.U. Leuven.
Debón, A., Montes, F., Sala, R., 2006. A comparison of models for dynamic life tables. Application to mortality data from the valencia region (Spain). Lifetime Data Anal. 12, 223–244.

Gaille, S., Sherris, M., 2011. Modeling mortality with common stochastic long-run trends. Geneva Pap. Risk Insur. - Issues Pract. 36 (4), 595–621.
Gonzalo, J., Granger, C., 1995. Estimation of common long-memory components in cointegrated systems. J. Bus. Econom. Statist. 13 (1), 27–35.
Hamilton, J.D., 1994. Time Series Analysis. Princeton University Press, Princeton.
Hayflick, L., 2004. "Anti-aging" is an oxymoron. J. Gerontol.: Biol. Sci. 59A (6), 573–578.
Hogg, R.V., McKean, J., Craig, A.T., 2014. Introduction to Mathematical Statistics. Pearson Education, Boston, Mass.
Holliday, R., 2004. The multiple and irreversible causes of aging. J. Gerontol.: Biol. Sci. 59A (6), 568–572.
Hunt, A., Blake, D., 2015. Modelling longevity bonds: Analysing the swiss re kortis bond. Insurance Math. Econom. 63, 12–29.
Jin, K., 2010. Modern biological theories of aging. Aging Dis. 1 (2), 72–74.
Johansen, S., 1988. Statistical analysis of cointegration vectors. J. Econom. Dynam. Control 12, 231–254.
Johansen, S., 1995. Likelihood-Based Inference in Cointegrated Vector Autoregressive Models. Oxford University Press, NewYork.
Jolliffe, I.T., 2002. Principal Component Analysis. Springer, New York.
Lazar, D., Denuit, M., 2009. A multivariate time series approach to projected life tables. Appl. Stoch. Models Bus. Ind. 25, 806–823.
Li, J.S.-H., Hardy, M.R., 2011. Measuring basis risk in longevity hedges. N. Am. Actuar. J. 15 (2), 177–200.
Li, H., Lu, Y., 2017. Coherent forecasting of mortality rates: a sparse vector-autoregression approach. Astin Bull. 47 (2), 563–600.
Lütkepohl, H., 2005. New Introduction to Multiple Time Series Analysis. Springer, Berlin.
Njenga, C., Sherris, M., 2011. Longevity risk and the econometric analysis of mortality trends and volatility. Asia-Pac. J. Risk Insur. 5 (2).
Olshansky, S.J., Hayflick, L., Carnes, B., 2002. Position statement on human aging. J. Gerontol.: Biol. Sci. 57A (8), B292–B297.
Salhi, Y., Loisel, S., 2017. Basis risk modelling: a co-integration based approach. J. Theor. Appl. Statist. 51 (1), 205–221.
Stock, J.H., Watson, M.W., 1988. Testing for common trends. Amer. Statist. Assoc. 83 (404), 1097–1107.
World Health Organization, 2016. WHO Mortality database. http://www.who.int/whosis/mort/download/en/index.html.
Yang, S., Wang, C.-W., 2013. Pricing and securitization of multi-country longevity risk with mortality dependence. Insurance Math. Econom. 52, 157–169.
Yang, S., Yue, J.C., Huang, H.-C., 2010. Modeling longevity risks using a principal component approach: A comparison with existing stochastic mortality models. Insurance Math. Econom. 46, 254–270.
Zhou, R., Wang, Y., Kaufhold, K., Siu-Han Li, J., Seng Tan, K., 2014. Modeling period effects in multi-population mortality models: Applications to solvency II. N. Am. Actuar. J. 18 (1), 150–167.